sequenced as described in ref. 18. Internal primers McytbseqF (5′-ATT GAC TAT GGC GAC CGC TTT T-3′) and McytbseqR (5′-GAA TAA AAT TCT CTG CGT CTC C-3′) for *cytB* were used in addition to PCR primers.

Primers to amplify the β-tubulin gene (intron and exon regions), designed on the basis of sequence data of *Montastraea faveolata*[19], were TubulinF (5′-GCA TGG GAA CGC TCC TTA TTT-3′) and TubulinR (5′-ACA TCT GTT GAG TGA GTT CTG-3′). They amplify a region corresponding to amino acid positions 144–299 within exon 4 of the human and *Drosophila* β-tubulin gene; the beginning of the intron corresponds to position 247, and the flanking exons have 99% amino acid similarity to the corresponding vertebrate sequence. Depending on the genus, one, two or three bands of about 600 bases, 1.0–1.5 kilobases (kb) and more than 2.0 kb were amplified by PCR with the above-described protocol. The size difference between bands was due to the length of the intron. Because most genera had a 1.0–1.5-kb band, this was used for phylogenetic analyses. *Diploastrea* and *Solenastrea* had only the 600-base-pair (bp) band, and no bands could be amplified for *Acanthastrea rotundoflora* and *Favites chinensis*; consequently these four taxa were not analysed for β-tubulin. Amplified fragments of the β-tubulin gene were separated by agarose electrophoresis, cloned with the pGEM-T System (Promega) and sequenced for both strands. At least five clones obtained from each of two independent PCRs were analysed. If only one sequence occurred more than once, this sequence was used in the phylogenetic analyses; otherwise the two most abundant sequences were used.

DNA phylogenetic analyses

Only the exon regions of the β-tubulin gene (444 bp) were analysed, because the intron was too variable for alignment. Phylogenetic analyses were performed with PAUP*[20]. DNA sequences of the entire *cytB* gene and the *COI* gene excluding the third codon position (total length 1,557 bases) were combined on the basis of nucleotide saturation analyses and the incongruence length difference test. Phylogenetic trees were constructed on the basis of neighbour-joining (NJ), maximum-parsimony (MP) and maximum-likelihood (ML) methods with the use of PAUP*. The NJ analysis was done with a two-parameter model[21]. In MP and ML analyses, heuristic searches with TBR branch swapping and 25 random additions of taxa were performed. For ML analysis, we used Modeltest[22] to find an appropriate model of evolution. For mitochondrial genes the K81uf model[23] with gamma parameter (G) and proportion of invariable positions (I) was chosen. For β-tubulin we chose the TrN model[24] with G and I. Bootstrapping was used to evaluate support for trees (1,000 replicates for NJ and MP; 300 bootstraps with the fast-stepwise heuristic search for ML).

Received 16 December 2003; accepted 13 January 2004; doi:10.1038/nature02339.

1. Veron, J. E. N. *Corals in Space and Time* (UNSW Press, Sydney, 1995).
2. Veron, J. E. N. *Corals of the World* (Australian Institute of Marine Science, Townsville, 2000).
3. Roberts, C. M. *et al.* Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* **295,** 1280–1284 (2002).
4. Budd, A. F. Diversity and extinction in the Cenozoic history of Caribbean reefs. *Coral Reefs* **19,** 25–35 (2000).
5. Romano, S. L. & Cairns, S. D. Molecular phylogenetic hypotheses for the evolution of scleractinian corals. *Bull. Mar. Sci.* **67,** 1043–1068 (2000).
6. Veron, J. E. N., Odorico, D. M., Chen, C. A. & Miller, D. J. Reassessing evolutionary relationships of scleractinian corals. *Coral Reefs* **15,** 1–9 (1996).
7. Chen, C. A., Wallace, C. C. & Wolstenholme, J. Analysis of the mitochondrial 12S rRNA gene supports a two-clade hypothesis of the evolutionary history of scleractinian corals. *Mol. Phyl. Evol* **23,** 137–149 (2002).
8. Cuif, J.-P., Lecointre, G., Perrin, C., Tillier, A. & Tillier, S. Patterns of septal biomineralization in Scleractinia compared with their 28S rRNA phylogeny: a dual approach for a new taxonomic framework. *Zool. Scripta* **32,** 459–473 (2003).
9. Vaughan, T. W. & Wells, J. W. Revision of the suborders, families, and genera of the Scleractinia. *Geol. Soc. Am. Spec. Pap.* **44,** 1–363 (1943).
10. Stolarski, J. & Roniewicz, E. Towards a new synthesis of evolutionary relationships and classification of Scleractinia. *J. Paleontol.* **75,** 1090–1108 (2001).
11. Wells, J. W. The recent solitary mussid scleractinian corals. *Zool. Meded.* **39,** 375–384 (1964).
12. Paulay, G. in *Life and Death of Coral Reefs* (ed. Birkeland, C.) 298–353 (Chapman & Hall, New York, 1997).
13. Wallace, C. C. Journey to the heart of the centre—origins of high marine faunal diversity in the central Indo-Pacific from the perspective of an acropologist. *Proc. 9th Int. Coral Reef Symp.* **1,** 33–39 (2002).
14. Shearer, T. L., van Oppen, M. J. H., Romano, S. L. & Wörheide, G. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol. Ecol.* **11,** 2475–2487 (2002).
15. Mace, G. M., Gittleman, J. L. & Purvis, A. Preserving the tree of life. *Science* **300,** 1707–1709 (2003).
16. Pandolfi, J. M. *et al.* Global trajectories of the long-term decline of coral reef ecosystems. *Science* **301,** 955–958 (2003).
17. Gardner, T. A., Côté, I. M., Gill, J. A., Grant, A. & Watkinson, A. R. Long-term region-wide declines in Caribbean corals. *Science* **301,** 958–960 (2003).
18. Williams, S. T., Knowlton, N., Weight, L. A. & Jara, J. A. Evidence for three major clades within the snapping shrimp genus *Alpheus* inferred from nuclear and mitochondrial gene sequence data. *Mol. Phyl. Evol.* **20,** 375–389 (2001).
19. Lopez, J. V. & Knowlton, N. Discrimination of sibling species in the *Montastraea annularis* complex using multiple genetic loci. *Proc. 8th Int. Coral Reef Symp.* **2,** 1613–1618 (1997).
20. Swofford, D. L. *PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods), Version 4.0b10* (Sinauer, Sunderland, MA, 2002).
21. Kimura, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16,** 111–120 (1980).
22. Posada, D. & Crandall, K. A. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **14,** 817–818 (1998).
23. Kimura, M. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl Acad. Sci. USA* **78,** 454–458 (1981).
24. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10,** 512–526 (1993).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank J. Jara, E. Gomez, M. Hatta and staff of the Palau International Coral Reef Center for their assistance in the field and laboratory, and J. Jackson and R. Grosberg for comments on the manuscript. Financial support came from the National Science Foundation, the Smithsonian Institution, the Scripps Institution of Oceanography, and the Conselho Nacional de Pesquisas (CNPq).

Authors' contributions H.F. performed the collections and genetic analyses, A.F.B. the morphological and palaeontological interpretations, C.A.C. the coral molecular systematics and collections in Taiwan, G.P. the coral systematics and collections in Palau, A.S.-C. the molecular systematics and collections in Brazil, K.I. the collections in Okinawa, and N.K. the coral systematics, financial/logistic support and manuscript preparation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.K. (nknowlton@ucsd.edu). DNA sequences are available in DDBJ (accession numbers AB117222–AB117388 and AB118246–AB118435).

Phylogenetic constraints and adaptation explain food-web structure

Marie-France Cattin[1], Louis-Félix Bersier[1,2], Carolin Banašek-Richter[1], Richard Baltensperger[3] & Jean-Pierre Gabriel[3]

[1]*Institut de Zoologie, University of Neuchâtel, Rue Emile-Argand 11, CP2, CH-2007 Neuchâtel, Switzerland*
[2]*Chair of Statistics, Department of Mathematics, Swiss Federal Institute of Technology, CH 1015 Lausanne, Switzerland*
[3]*Department of Mathematics, Université de Fribourg, Pérolles, CH 1700 Fribourg, Switzerland*

Food webs are descriptions of who eats whom in an ecosystem. Although extremely complex and variable, their structure possesses basic regularities[1–6]. A fascinating question is to find a simple model capturing the underlying processes behind these repeatable patterns. Until now, two models have been devised for the description of trophic interactions within a natural community[7,8]. Both are essentially based on the concept of ecological niche, with the consumers organized along a single niche dimension; for example, prey size[8,9]. Unfortunately, they fail to describe adequately recent and high-quality data. Here, we propose a new model built on the hypothesis that any species' diet is the consequence of phylogenetic constraints and adaptation. Simple rules incorporating both concepts yield food webs whose structure is very close to real data. Consumers are organized in groups forming a nested hierarchy, which better reflects the complexity and multidimensionality of most natural systems.

A central issue in ecology is to uncover the basic determinants of the distribution of trophic interactions among the members of natural communities[9]. The architecture of interactions affects the stability properties of dynamical models of food webs[2,10,11]. Therefore, a full understanding of dynamic ecosystems cannot be achieved at the economy of assuming a static structure of food webs, as was the case in the pioneering works on stability and complexity that considered interactions to be random[12,13]. It has been shown unambiguously that real food webs are different from randomly connected networks, and that such a null-model cannot account for the observed properties of the highest-quality food webs available[6,8]. The structural models of trophic interactions proposed

so far are the cascade model[7] and the niche model[8]. The former assumes that species are ranked from 1 to S (total number of species), and that a consumer preys only on species of lower rank, with probability $P = 2SC/(S-1)$ (directed connectance $C = L/S^2$, where L is the total number of trophic links). The niche model orders the species according to a randomly drawn 'niche value', n_i ($0 \leq n_i \leq 1$). A consumer eats all species falling into a range whose centre c_i is randomly chosen, with $c_i < n_i$. This restricts diets to being contiguous (Fig. 1a). Contiguity reflects the ecological assumption that diets can be arranged along one niche dimension.

Whereas the cascade model fails to describe real data, the niche model closely predicts properties of recent high-quality food webs[8]. However, it does not totally account for the observed patterns. The model is known to produce only interval food webs (Fig. 2a, b). These interval webs possess the intriguing characteristic that the feeding relationships between consumers can be represented in one dimension[1,14], thought to correspond to a body size hierarchy[9]. But, non-interval food webs characterize recent high-quality data[8]. A more fundamental problem is that a key assumption of the niche model, contiguity of diets, is not observed in real food webs. By constraining consumers to eat all species in a range, the niche model entails no gap in diets for a suitable ordering of the prey (no gap in columns of Fig. 1a). In a food-web matrix, it is often computationally impracticable to find the prey sequence that would yield the minimum number of gaps. However, an irreducible gap can occur with at least three non-monophagous consumers (a triplet; see Fig. 2c). Consequently, we propose a new property of food-web structure—the level of diet discontinuity, D_{diet}—defined as the proportion of triplets with an irreducible gap over the number of possible triplets. The niche model predicts D_{diet} to equal 0. In seven of the largest and highest-quality empirical food webs (see Methods), the mean of D_{diet} is 0.18, a figure significantly different from 0 ($P = 0.006$, one-tailed, one-sample t-test; Table 1).

These difficulties prompted us to take an evolutionary perspective of the basic ecological determinants of trophic interactions. Typically, a consumer's diet is constrained by its phylogenetic origin[15]. For example, all warblers of the *Phylloscopus* genus possess a beak suited to preying on insects; all locusts of the Acrididae family have mouthparts and an alimentary tract suited for plant material. To substantiate the connection between phylogenetic origin and diet, we compared the matrices of trophic and taxonomic similarity for five empirical food webs by Mantel tests[16] (see Methods). There was a strong relationship (all P-values <0.001), indicating that the distribution of trophic interactions is indeed related to the structure imposed by the phylogeny of the species forming the community.

However, the variability seen in the trophic relations between species cannot be explained fully by phylogenetic constraints. Species have to adapt to varying environments in order to survive, diverging from close relatives in their behaviour, and possibly innovating by using new food sources. Within this framework for example, the diet of New Guinea's fruit pigeons[17] may be seen as constrained by the type of beak and digestive system characteristic of Columbiformes, and the partitioning of fruit size by adaptation to the biotic environment. Dietary shifts also occur frequently, especially in higher animals that are known to be opportunistic. In accordance with the macroevolutionary theory[18], we feel that both phylogenetic constraints and adaptation are the essential determi-

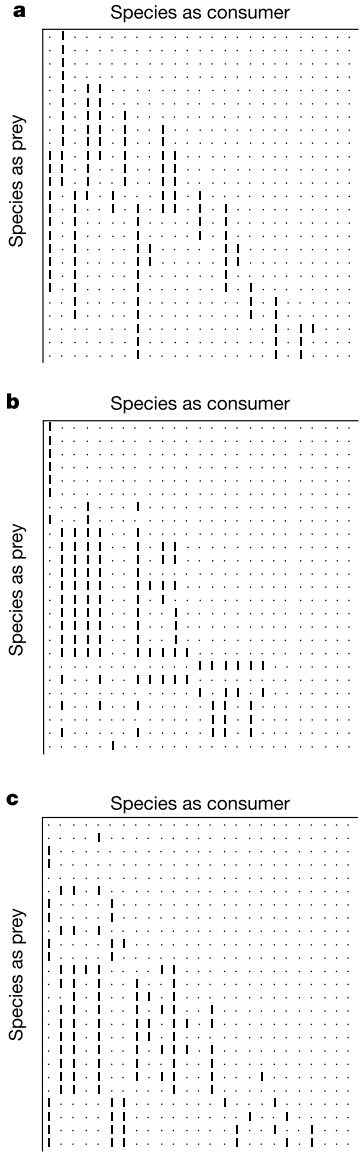

Figure 1 Comparison of one simulation of the niche model (**a**) and of the nested hierarchy model (**c**), with respect to a real food web (**b**); Bridge Brook Lake[4]. The food-web matrix depicts the trophic relationships, where a vertical mark indicates that the species in the column consumes the species in the row (columns and rows contain the same species).

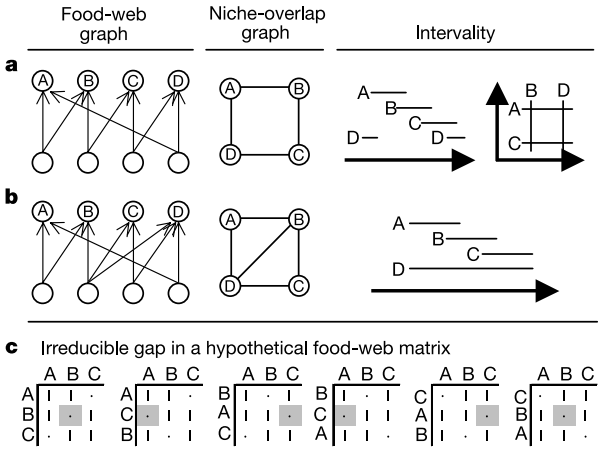

Figure 2 Hypothetical food webs illustrating chordless cycles and intervality, and irreducible gaps. **a**, **b**, Food-web graph. Circles are species and arrows are flows; A–D are consumers. The niche-overlap graph is an undirected graph where consumers sharing prey are connected. Under 'intervality', graphs depict consumers as segments; segments overlap if consumers share prey. **a**, A chordless and non-interval food web. The niche-overlap graph is chordless because no edge links A to C and/or B to D; the food web is non-interval (consumer D (broken line) cannot be placed in one dimension, because it overlaps with A and C, but not with B; thus, two dimensions are needed). **b**, A non-chordless and interval food web. **c**, Irreducible gaps.

Table 1 **Comparison of properties for seven food webs**

Property	Skipwith Pond			Little Rock Lake			Bridge Brook Lake			Chesapeake Bay			Ythan Estuary			Coachella Desert			St Martin Island		
T	0.06	[0.04]	0.08	0.05	[0]	0.03	0.10	[0]	0.10	0.14	[0.28]	0.15	0.08	[0.38]	0.07	0.06	[0]	0.08	0.09	[0.17]	0.08
I	0.85	[0.92]	0.82	0.85	[0.87]	0.88	0.73	[0.68]	0.74	0.60	[0.62]	0.59	0.74	[0.53]	0.77	0.86	[0.90]	0.83	0.75	[0.69]	0.77
B	0.09	[0.04]	0.10	0.10	[0.13]	0.08	0.17	[0.32]	0.16	0.26	[0.10]	0.26	0.18	[0.09]	0.17	0.09	[0.10]	0.08	0.17	[0.14]	0.15
Gen_{sd}	0.76	[0.92]	0.83	1.06	[1.40]	1.08	0.98	[1.09]	1.03	1.12	[0.71]	1.17	1.17	[1.16]	1.18	0.81	[0.73]	0.83	1.06	[1.02]	1.09
Vul_{Sd}	0.55	[0.54]	0.59	0.59	[0.57]	0.59	0.63	[0.61]	0.72	0.73	[1.03]	0.80	0.66	[1.40]	0.71	0.53	[0.60]	0.60	0.64	[0.78]	0.71
M_{sim}	0.73	[0.76]	0.67	0.69	[0.76]	0.54	0.61	[0.71]	0.56	0.48	[0.34]	0.48	0.58	[0.50]	0.50	0.74	[0.72]	0.67	0.62	[0.54]	0.53
Ch_{mean}	6.23	[4.81]	5.38		–		4.86	[2.55]	4.30	3.28	[2.77]	3.30	6.39	[4.89]	6.94	6.95	[7.18]	5.78	5.76	[4.20]	5.62
Ch_{sd}	1.56	[1.32]	1.58		–		1.51	[0.76]	1.52	1.36	[1.14]	1.42	1.89	[1.50]	2.17	1.68	[1.89]	1.64	1.65	[1.30]	1.82
Ch_{log}	4.21	[3.52]	3.70		–		3.20	[2.62]	2.90	2.35	[2.11]	2.32	4.39	[3.99]	4.38	4.62	[4.90]	4.04	3.96	[3.52]	3.71
Lo	0.25	[0.12]	0.40		–		0.05	[0]	0.13	0.00	[0.24]	0.02	0.01	[0]	0.06	0.26	[0]	0.42	0.03	[0]	0.11
Can_{sp}	0.43	[0.32]	0.10	0.14	[0.15]	0.03	0.21	[0.12]	0.03	0.08	[0]	0.00	0.07	[0.04]	0.01	0.44	[0.66]	0.09	0.13	[0]	0.02
O	0.77	[0.6]	0.76		–		0.61	[0.36]	0.60	0.42	[0.38]	0.41	0.61	[0.53]	0.61	0.75	[0.79]	0.78	0.63	[0.60]	0.63
Cy_4	0	[0]	0	0	[639]	626	0	[0]	0	0	[0]	0	0	[206]	18	0	[36]	1	0	[226]	6
D_{diet}	0	[0.35]	0.16	0	[0.08]	0.09	0	[0.004]	0.09	0	[0.08]	0.02	0	[0.32]	0.02	0	[0.28]	0.17	0	[0.17]	0.07

For each food web, empirical values are in brackets, means (medians for Cy_4) for the niche model are on the left, and those for the nested hierarchy model on the right. See Methods for a description of the properties of the food webs. Little Rock Lake food web is too large for Ch_{mean}, Ch_{sd}, Ch_{log}, Lo and O to be computed in a reasonable length of time.

nants of the distribution of trophic interactions in any community.

We devised simple rules to generate food-web matrices according to both concepts. A food-web matrix is an S-by-S matrix with species as prey in rows and the same list of species as consumers in columns, with a '1' indicating that the species in the column consumes the species in the row (see Fig. 1). As with the cascade and niche models, parameters S and L are needed. We followed ref. 8 to set the number of links per consumer, which requires species to be ordered according to their 'niche value'; species with the smallest niche value tend to have the smallest number of prey (see Methods). The trophic links are attributed to consumers in a two-stage process, starting with the smallest niche value. In stage one, prey species of consumer i is randomly chosen among species with rank $<$i. Depending on this randomly chosen prey j, two cases are possible: (1) prey j has no consumer and therefore the next prey of consumer i will again be randomly attributed (with rank of prey $<$i); (2) prey j already has one or more consumers and therefore consumer i joins the group of species j's consumers (that is, all consumers sharing at least one prey, with at least one consumer of this group feeding on j), and the next prey of consumer i is then randomly chosen among the set of prey of this group. However, if the number of prey in the group is too small for choosing all remaining prey of consumer i, the remaining prey are again (randomly) chosen among prey without consumers (with rank of prey $<$i). The second stage is needed if prey still cannot be attributed; remaining prey of consumers for which prey could not be attributed in stage 1 are randomly chosen (prey species can have rank $\geq$i).

By creating groups of consumers, stage one (2) expresses the part in food-web organization that is determined by phylogenetic constraints. Links attributed to species free of consumers, and links distributed randomly in the second stage, render the adaptation of consumers to new prey. In forcing consumers to form various trophic groups, our 'nested-hierarchy model' escapes the one-dimensional nature of former models, and better reflects the kind of hierarchies emanating from the phylogenetic structure of the community.

We analysed 12 properties of food webs generated by the niche model as opposed to food webs generated by our model (see Methods). The results of 1,000 simulations of each model for webs with $S = 20$–100 ($S = 20$–50 for computer-intensive properties) and $C = 0.11$ (ref. 19) show that both models yield quite similar results (Fig. 3). When compared with the seven reference food webs, we observe that the niche model performs slightly better for five properties (Gen_{sd}, M_{sim}, Lo, Can_{sp}, O), that our model is better for four properties (B, Vul_{sd}, Ch_{mean}, Ch_{log}), and that both models are equal for T, I and Ch_{sd} (Wilcoxon paired-sample tests, Table 1; see Methods for definitions). In sum, our model performs as well as the niche model—itself outperforming the cascade model by one order of magnitude—in the prediction of these standard food-web descriptors.

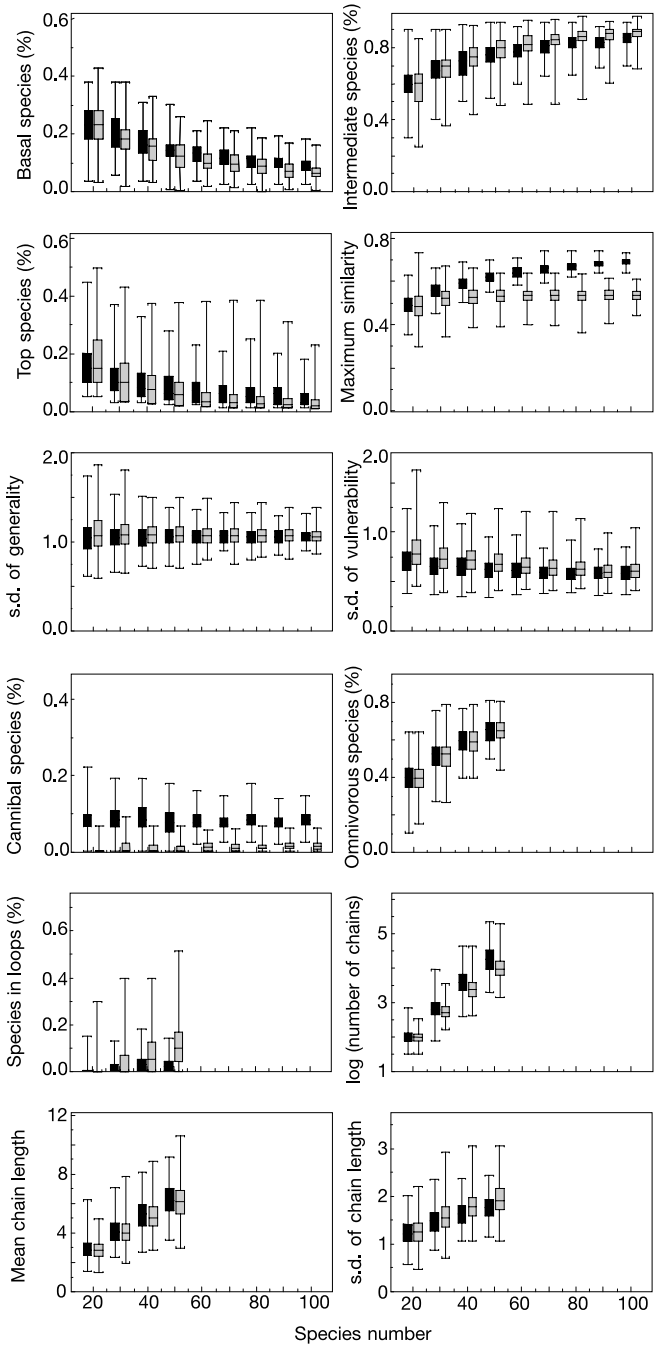

Figure 3 Comparison of 12 properties (see Methods) of food webs generated by the niche model (black boxes) and the nested-hierarchy model (grey boxes).

The nested-hierarchy model surpasses the niche model for two additional properties: D_{diet} and the number of chordless cycles, Cy_4 (Table 1). The latter is closely linked to intervality, as interval food webs have no chordless cycles[2] (Fig. 2); Cy_4 may thus be thought of as the degree of departure from intervality (that is, from the possibility of ordering the consumers along a single dimension). Finally, the nested-hierarchy model generates food-web matrices whose global structure is close to the observed ones (Fig. 1c).

There are surely other conceivable ways in which to include phylogenetic constraints and adaptation; the one we devised was for us the most intuitive and most parsimonious. The fact that such simple rules are able to faithfully reproduce objects as complex as food webs is a strong hint at their usefulness; for example, as a roadmap for the interactions in dynamical models. What we perceive to be of higher importance than details of model construction are the processes behind the nested-hierarchy model. We have shown how phylogeny is intimately linked to trophic structure in natural communities, and how, once included in a model, it closely predicts observed patterns. When considering trophic links, body size is without doubt another constraint that will limit possible interactions[20]. Yet body size varies widely within trophic groups, and the size distributions of different trophic groups can overlap extensively—simply think of the herbivores and carnivores in African savannahs. As a consequence, body size is of secondary importance in explaining food-web structure when compared with phylogeny.

Compared with the niche and the cascade models, we impose a sequence in the attribution of links, because a consumer's diet will depend on the species already present. Accordingly, prey are first attributed to primary consumers, as expected in the course of community development. This sequential process is inspired by the niche-hierarchy model[21], an assembly rule stating that species joining a community will be successful only if they compete within single guilds. For example, it would not be possible for a bird feeding on fruits and insects to enter a community if guilds of frugivores and insectivores are already present. The niche-hierarchy model will also generate nested hierarchies with groups defined by their feeding habits. In our model, we propose that this grouping is constrained by phylogeny, but we permit species to break the strict rule of the niche-hierarchy model by allowing consumers to prey on species outside their group. The niche-hierarchy model has recently been extended to explain patterns of species abundances[22]. An intriguing perspective would be to combine the niche-hierarchy with our nested-hierarchy model. It will offer a framework to get a sharper understanding of quantitative food webs, including the relationship between abundance, body size and trophic structure[20]. □

Methods

Data set

The food webs considered here and in ref. 8 are the following (N = total number of taxa, S = total number of trophic species, C = directed connectance): Skipwith pond[23] ($N = 35$, $S = 25$, $C = 0.32$); Chesapeake Bay[24] ($N = 33$, $S = 31$, $C = 0.072$); Ythan Estuary[25] ($N = 92$, $S = 79$, $C = 0.061$); Coachella Valley[26] ($N = 30$, $S = 29$, $C = 0.31$); St Martin Island[27] ($N = 44$, $S = 42$, $C = 0.12$); Little Rock Lake[28] ($N = 180$, $S = 94$, $C = 0.12$); and Bridge Brook Lake[4] ($N = 75$, $S = 25$, $C = 0.17$). A trophic species is defined as a set of taxa sharing the same prey and consumers[8].

Mantel test

Trophic similarity between two species i and j was measured with the index of Jaccard[29]; that is, as the number of prey and of consumers shared by i and j divided by the pair's total number of prey and consumers. To measure taxonomic similarity, we first ascribed to each taxon (for economy, we use the word 'species' in the text, but prey and consumers are often described at a coarser taxonomic level) its taxonomic membership. We used ten levels: (1) kingdom, (2) superphylum, (3) phylum, (4) subphylum, (5) class, (6) subclass/superorder, (7) order, (8) suborder/superfamily, (9) family, (10) genus. Taxonomic similarity between two taxa, i and j, was measured as the value of the most precise common taxonomic level (for example, ten for two species of the same genus) divided by one plus the value of the most detailed level of any of both taxa. We performed a Mantel test[16] to compare the matrices of trophic and taxonomic similarity of Skipwith pond, Chesapeake Bay, Ythan Estuary, Coachella Valley and St Martin Island. Real taxa and not trophic species were used. Little Rock Lake and Bridge Brook Lake were not considered because taxonomic information was used to infer trophic interactions.

Model

To determine the number of links per consumer l_i, we assigned each species a niche value n_i uniformly drawn from [0,1] and sorted species according to their n_i in ascending order. Each l_i was obtained by multiplying n_i by a value drawn from [0,1], using a β-distribution with expected value $2C$ and with $\alpha = 1$ (see ref. 8). To obtain the desired L, each consumer's l_i was divided by the sum of l_i over all species and multiplied by L. If the resulting l_i exceeded $S - 1$, we arbitrarily reduced it to $S - 1$. To ensure at least one basal species, the species with the smallest n_i had no prey. For computational reasons, we imposed the presence of at least one top species (the species with the largest n_i has no consumer). We excluded webs with connectance less than 97% of the desired C, with disconnected species, or with species sharing the same consumers and prey.

Properties

The following 14 properties were calculated for each web: B, I, and T, the proportions of basal (without prey), intermediate (with both predators and prey) and top (without predators) species, respectively; Gen_{sd} and Vul_{sd}, the standard deviations (s.d.) of generality and vulnerability[8], which measure the variation in prey and predator numbers per species, respectively; M_{sim}, the mean maximum similarity of a web, which is the average of all species' largest value for the trophic similarity; Can_{sp}, the proportion of cannibalistic species; O, the proportion of omnivores (that is, species that consume at least two species and have food chains of different lengths); $Ch_{\log}$, the log of the number of food chains (a food chain is a linked path between any species to a basal species); Ch_{mean} and Ch_{sd}, the mean and the s.d. of food chain lengths; Lo, the proportion of species involved in loops other than cannibalistic loops (that is, parts of food chains that include the same species twice). Links involved in loops are ignored for the computation of $Ch_{\log}$, Ch_{mean}, Ch_{sd} and O. The final two properties are Cy_4, the number of chordless cycles of length four (see Fig. 2), and D_{diet}, the level of diet discontinuity, which measures the proportion of triplets of consumers whose prey cannot be ordered so that the three diets are fully contiguous.

Received 27 October 2003; accepted 12 January 2004; doi:10.1038/nature02327.

1. Cohen, J. E. *Food Webs and Niche Space* (Princeton Univ. Press, Princeton, New Jersey, 1978).
2. Sugihara, G. *Niche Hierarchy: Structure Assembly and Organization in Natural Communities.* PhD Thesis, Princeton Univ., Princeton (1982).
3. Cohen, J. E., Briand, F. & Newman, C. N. *Community Food Webs: Data and Theory* (Springer, Berlin, 1990).
4. Havens, K. E. Scale and Structure in Natural Food Webs. *Science* **257,** 1107–1109 (1992).
5. Bersier, L. F. & Sugihara, G. Scaling regions for food web properties. *Proc. Natl Acad. Sci. USA* **94,** 1247–1251 (1997).
6. Solow, A. R. & Beet, A. R. On lumping species in food webs. *Ecology* **79,** 2013–2018 (1998).
7. Cohen, J. E. & Newman, C. M. A stochastic theory of community food webs I. Models and aggregated data. *Proc. R. Soc. Lond. B* **224,** 449–461 (1985).
8. Williams, R. J. & Martinez, N. D. Simple rules yield complex food webs. *Nature* **404,** 180–183 (2000).
9. Lawton, J. H. & Warren, P. H. Static and dynamic explanations for patterns in food webs. *Trends Ecol. Evol.* **3,** 242–245 (1988).
10. DeAngelis, D. L. Stability and connectance in food web models. *Ecology* **56,** 238–243 (1975).
11. Kondoh, M. Foraging adaptation and the relationship between food-web complexity and stability. *Science* **299,** 1388–1391 (2003).
12. Gardner, M. & Ashby, W. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature* **228,** 784 (1970).
13. May, R. M. *Stability and Complexity of Model Ecosystems* (Princeton Univ. Press, Princeton, New Jersey, 1974).
14. Pimm, S. L. *Food Webs* (Chapman and Hall, London, 1982).
15. Cousins, S. H. The trophic continuum in marine ecosystems: Structure and equations for a predictive model. *Can. J. Fish. Aquat. Sci.* **213,** 76–93 (1985).
16. Legendre, P. & Legendre, L. *Numerical Ecology* (Elsevier, Amsterdam, 1998).
17. Diamond, J. M. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 342–444 (Belknap/Harvard Univ. Press, Cambridge, Massachusetts, 1975).
18. Price, P. W. *Macroevolutionary Theory on Macroecological Patterns* (Cambridge Univ. Press, Cambridge, 2003).
19. Martinez, N. D. Constant connectance in community food webs. *Am. Nat.* **139,** 1208–1218 (1992).
20. Cohen, J. E., Jonsson, T. & Carpenter, S. R. Ecological community description using the food web, species abundance, and body size. *Proc. Natl Acad. Sci. USA* **100,** 1781–1786 (2003).
21. Sugihara, G. in *Population Biology. Proceedings of Symposia in Applied Mathematics* (ed. Levin, S. A.) 83–101 (American Mathematical Society, Providence, Rhode Island, 1984).
22. Sugihara, G., Bersier, L. F., Southwood, T. R. E., Pimm, S. L. & May, R. M. Predicted correspondence between species abundances and dendrograms of niche similarities. *Proc. Natl Acad. Sci. USA* **100,** 5246–5251 (2003).
23. Warren, P. H. Spatial and temporal variation in the structure of a freshwater food web. *Oikos* **55,** 299–311 (1989).
24. Baird, D. & Ulanowicz, R. E. The seasonal dynamics of the Chesapeake bay ecosystem. *Ecol. Monogr.* **59,** 329–364 (1989).
25. Hall, S. J. & Raffaelli, D. Food-web patterns: lessons from a species-rich web. *J. Anim. Ecol.* **60,** 823–842 (1991).
26. Polis, G. A. Complex trophic interactions in deserts: an empirical critique of food-web theory. *Am. Nat.* **138,** 123–155 (1991).
27. Goldwasser, L. & Roughgarden, J. Construction and analysis of a large Caribbean food web. *Ecology* **74,** 1216–1233 (1993).
28. Martinez, N. D. Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.* **61,** 367–392 (1991).

29. Jaccard, P. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaudoise Sci. Nat.* **44,** 223–270 (1908).

Acknowledgements We thank J. Bascompte, S. Cousins, S. Hubbell, R. Naisbit and P. Warren for useful comments. This work was funded by the Swiss National Science Foundation, the Novartis Foundation, and partly by the National Center of Competence in Research 'Plant Survival'.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.-F.B. (louis-felix.bersier@unine.ch).

Global organization of metabolic fluxes in the bacterium *Escherichia coli*

E. Almaas[1], B. Kovács[1,2], T. Vicsek[2], Z. N. Oltvai[3] & A.-L. Barabási[1]

[1]*Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA*
[2]*Biological Physics Department and Research Group of HAS, Eötvös University, H-1117 Budapest, Hungary*
[3]*Department of Pathology, Northwestern University, Chicago, Illinois 60611, USA*

Cellular metabolism, the integrated interconversion of thousands of metabolic substrates through enzyme-catalysed biochemical reactions, is the most investigated complex intracellular web of molecular interactions. Although the topological organization of individual reactions into metabolic networks is well understood[1–4], the principles that govern their global functional use under different growth conditions raise many unanswered questions[5–7]. By implementing a flux balance analysis[8–12] of the metabolism of *Escherichia coli* strain MG1655, here we show that network use is highly uneven. Whereas most metabolic reactions have low fluxes, the overall activity of the metabolism is dominated by several reactions with very high fluxes. *E. coli* responds to changes in growth conditions by reorganizing the rates of selected fluxes predominantly within this high-flux backbone. This behaviour probably represents a universal feature of metabolic activity in all cells, with potential implications for metabolic engineering.

To identify the interplay between the underlying topology[1–3] of the metabolic network of *E. coli* K12 MG1655 and its functional organization, we focus on the global features of potentially achievable flux states in this model organism with a fully sequenced and annotated genome[13,14]. In accordance with flux balance analysis (FBA)[8–12], we first identified the solution space (that is, all possible flux states under a given condition) by using constraints imposed by the conservation of mass and the stoichiometry of the reaction system for the reconstructed *E. coli* metabolic network[8–12]. Assuming that cellular metabolism is in a steady state and optimized for the maximal growth rate, FBA allows us to calculate the flux for each reaction using linear optimization[8–11], which provides a measure of the relative activity of each reaction.

As previously shown[8–10], the steady state and optimality approximations offer experimentally verifiable predictions on the flux states of the cell. Under any condition, however, there are also expected differences, some stemming from the fact that there are transient effects and that the growth rate is not always exactly optimal[12]. A marked feature of the obtained flux distribution is its overall inhomogeneity: reactions with fluxes spanning several orders of magnitude coexist under the same conditions. For example, under optimal growth conditions in a glutamate-rich culture, the dimensionless flux of the succinyl coenzyme A synthetase reaction is 0.185, whereas the flux of the aspartate oxidase reaction is four orders of magnitude smaller, with a value of 2.2×10^{-5} in dimensionless units (the flux vector is normalized to unity).

To characterize the coexistence of such widely different flux values, we plot the flux distribution for active (non-zero flux) reactions of *E. coli* grown in a glutamate- or succinate-rich substrate (Fig. 1a). The distribution is best-fitted with a power law with a small flux constant, indicating that the probability that a reaction has flux ν follows $P(\nu) \propto (\nu + \nu_0)^{-\alpha}$, where the constant is $\nu_0 = 0.0003$ and the flux exponent has the value $\alpha = 1.5$. The observed power law is also consistent with published experimental data. Indeed, the flux distribution obtained from the measured fluxes of the central metabolism of *E. coli*[15] is best-fitted with the power-law form (Fig. 1d). As the central metabolism is characterized by high fluxes, the small-flux saturation seen in Fig. 1a is absent from these data.

To examine whether the observed flux distribution is independent of the environmental conditions, we mimicked the influence of various growth conditions by randomly choosing 10%, 50% or 80% of the 96 potential substrates that *E. coli* can consume (the input substrates are listed in Supplementary Table S2). After optimizing the growth rate, we find that the power-law distribution of metabolic fluxes is independent of the external conditions (Fig. 1b). As the metabolic activity of *E. coli* frequently deviates from the optimal growth state under variable growth conditions[10,12], we inspected whether the wide flux distribution is also present in non-optimal

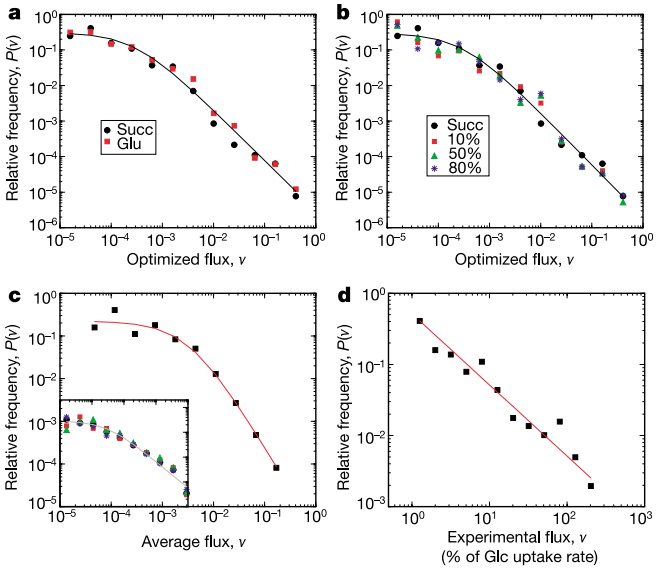

Figure 1 Characterizing the overall flux organization of the *E. coli* metabolic network. **a**, Flux distribution for optimized biomass production on succinate (black) and glutamate (red) substrates. Solid line corresponds to the power-law fit $P(\nu) \propto (\nu + \nu_0)^{-\alpha}$, with $\nu_0 = 0.0003$ and $\alpha = 1.5$. **b**, Flux distribution for optimized biomass on succinate (black) substrate with an additional 10% (red), 50% (green) and 80% (blue) randomly chosen subsets of the 96 input channels turned on. The flux distribution was averaged over 5,000 independent random choices of uptake metabolites. The resulting flux distribution can be fitted (solid line) by a power law with parameters $\nu_0 = 0.0004$ and $\alpha = 1.5$. **c**, Flux distribution from the non-optimized hit-and-run sampling method[16,17] of the *E. coli* solution space. The solid line is the best fit, with $\nu_0 = 0.003$ and $\alpha = 2$. Inset shows the flux distribution in four randomly chosen sample points. **d**, The distribution of experimentally determined fluxes (see ref. 15) from the central metabolism of *E. coli* shows power-law behaviour, with a best fit to $P(\nu) \propto \nu^{-\alpha}$ with $\alpha = 1$.